

Purdue University
Purdue e-Pubs

Libraries Research Publications

9-5-2007

Amelia Earhart's Missing Electra (Turboprop transports): controlled vocabulary in Contentdm

Janet V. Addison

Purdue University, jaddison@purdue.edu

Follow this and additional works at: http://docs.lib.purdue.edu/lib_research

Addison, Janet V., "Amelia Earhart's Missing Electra (Turboprop transports): controlled vocabulary in Contentdm" (2007). *Libraries Research Publications*. Paper 79.

http://docs.lib.purdue.edu/lib_research/79

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Amelia Earhart's missing Electra (Turboprop transports)

Controlled vocabulary in
Contentdm

Janet V. Addison
Society of American Archivists meeting
August 30, 2007
jaddison@purdue.edu

Outline

- Background to the digital collection
- Statement of immediate issue; overview of surrounding issues of controlled vocabulary
- Some foundations of controlled vocabulary
- Desirables
- Wrap-up and conclusions

George Palmer Putnam Collection of Amelia Earhart Papers

- Gifts from George Palmer Putnam (Amelia's husband), 1940; and Sally Putnam Chapman (George Putnam's granddaughter), 2002
- 16.5 cubic feet
- Correspondence, photographs, creative writings, printed material, scrapbooks, personal belongings, and ephemera relating to the life and career of aviator Amelia Earhart
- Bulk 1928-1937

George Palmer Putnam Collection of Amelia Earhart Papers : digital

- 2 digitization projects 1998; 2005
- 1998 project using home-grown system
- 2005 project using Contentdm
- Thesaurus used: Thesaurus for Graphic Materials (TGM)

Choosing Contentdm: cataloger's viewpoint

- Comes pre-loaded with Thesaurus for Graphic Materials (TGM) which we had already used in previous project
- LCSH and MeSH can be downloaded through Contentdm support page
- Ability to create cross-references
- OAI compliant: Metadata fields map to simple Dublin Core
- Words in item-level metadata are hyperlinked for additional searching

Electra (Turboprop transports)

- Library of Congress Subject Heading (LCSH)
- LCSH terms mixed in with TGM terms



CATALOGER'S DESKTOP

Find

[Advanced Search](#) [Search Tips & Search History](#)[Browse](#) [Document](#) [Results](#) [Doc/Results](#) [Search](#) [Preferences](#) [What's New](#) [Help](#) [Bookmarks/Notes](#) [Feedback](#) [Logout](#)[★ Start Over](#) [🖨 Print](#) [↶ Backtrack](#) [⏪ Prev Section](#) [⏩ Next Section](#) [🧼 Clear Search](#) [🔖 Bookmark/Note](#)

Airplanes

--[nationality]--[country or state]--[city]

Public Note Includes commercial, private, and military airplanes.

Broader Term [Aircraft](#)

Narrower Term [Airtankers](#)

[Biplanes](#)

[Bombers](#)

[Convertiplanes](#)

[Fighter planes](#)

[Gliders \(Aircraft\)](#)

[Monoplanes](#)

[Seaplanes](#)

[Supersonic planes](#)

[Transport planes](#)

Related Term [Air mail service](#)

[Airline industry](#)

[Airplane cockpits](#)

[Airplane equipment](#)

Electra (Turboprop transports)

Consistent complaint from users

... it seems your website almost always refers to the Lockheed Electra turboprop that was developed much later in the 50's, rather than the original Lockheed Electra that Amelia flew in the 1930's which was not a turboprop.

Electra (Greek mythology) and Electra (Carmen)





Lockheed Electra became new authorized subject heading

Subsequent effects

- Need cross reference from Electra (Turboprop transports)
- Cross-reference and heading conflicts with LCSH and with our OPAC entry
- Mixing TGM and LCSH with no differentiating labels

Final chart of decision points for AE controlled vocabulary

- Persons: Library of Congress Name Authority File (LCNAF)
- Geographic Names: Library of Congress Subject Headings (LCSH)
- Topics: (TGM I) supplemented by LCSH
- Genres: Thesaurus of Graphic Materials II (TGM II)
- **Local terms**



| | |
|---------------------|---|
| Old ID Number | XV.C.5.a |
| New ID Number | b12f9i29 |
| Title | Amelia Earhart walking in front of her Lockheed Electra |
| Description | Amelia Earhart walking in front of her Lockheed Electra airplane, ca. 1930s [One original and two poor duplicate photos] |
| Repository | Purdue University Libraries, Archives and Special Collections |
| Date of Original | 1930-1937 |
| Type | image |
| Format | JP2 |
| Extent of Original | 1 photograph : b&w ; 20 x 25 cm. |
| Orientation | Landscape |
| Collection | George Palmer Putnam Collection of Amelia Earhart Papers |
| Subjects | Earhart, Amelia, 1897-1937--Photographs Lockheed Electra Walking Photographic prints |
| Collection Web Site | http://www.lib.purdue.edu/earhart |
| Capture Device | Epson Expression 10000XL Photo Scanner |
| Rights | http://www.lib.purdue.edu/cdm4/item_viewer.php?CISOROOT=/earhart&CISOPTR=351&CISOBX=1&REC=15 |

Controlled vocabulary: initial decision points

- Is a controlled vocabulary needed
- Which thesaurus will be used
- One thesaurus or multiple thesauri
- Not just topical terms, but names, places
- Consider form/genre headings also

A controlled vocabulary: For all intents and purposes

- Increase number of access points
- Allow consistency
- Gather like items
- Promote precision of searching

Controlled vocabulary: quote

- “The best practice for subjects in shareable metadata is to use relevant controlled vocabularies consistently and explicitly. The controlled vocabularies chosen (there may be more than one used) should be relevant to the resource and known to the community to which the resources described would hold the most interest.”

Source: *Best Practices for OAI data provider implementations and shareable metadata* (wiki)

Especially helpful for visual materials

- Subject searching involves searching a standardized set of vocabulary terms
- Using the same terminology throughout a database creates **consistency** and **precision**, thus helping the user find relevant information no matter what terminology is used elsewhere in the record

Thesaurus choice

- Specificity and stability of the terms
- Content of the collection
- Homogeneous collection (e.g., all photographs or heterogeneous (e.g., photographs, realia, letters, video, etc.)
- Online search functionality
- Chief implementers/maintainers of the vocabulary
- Users performing the searching

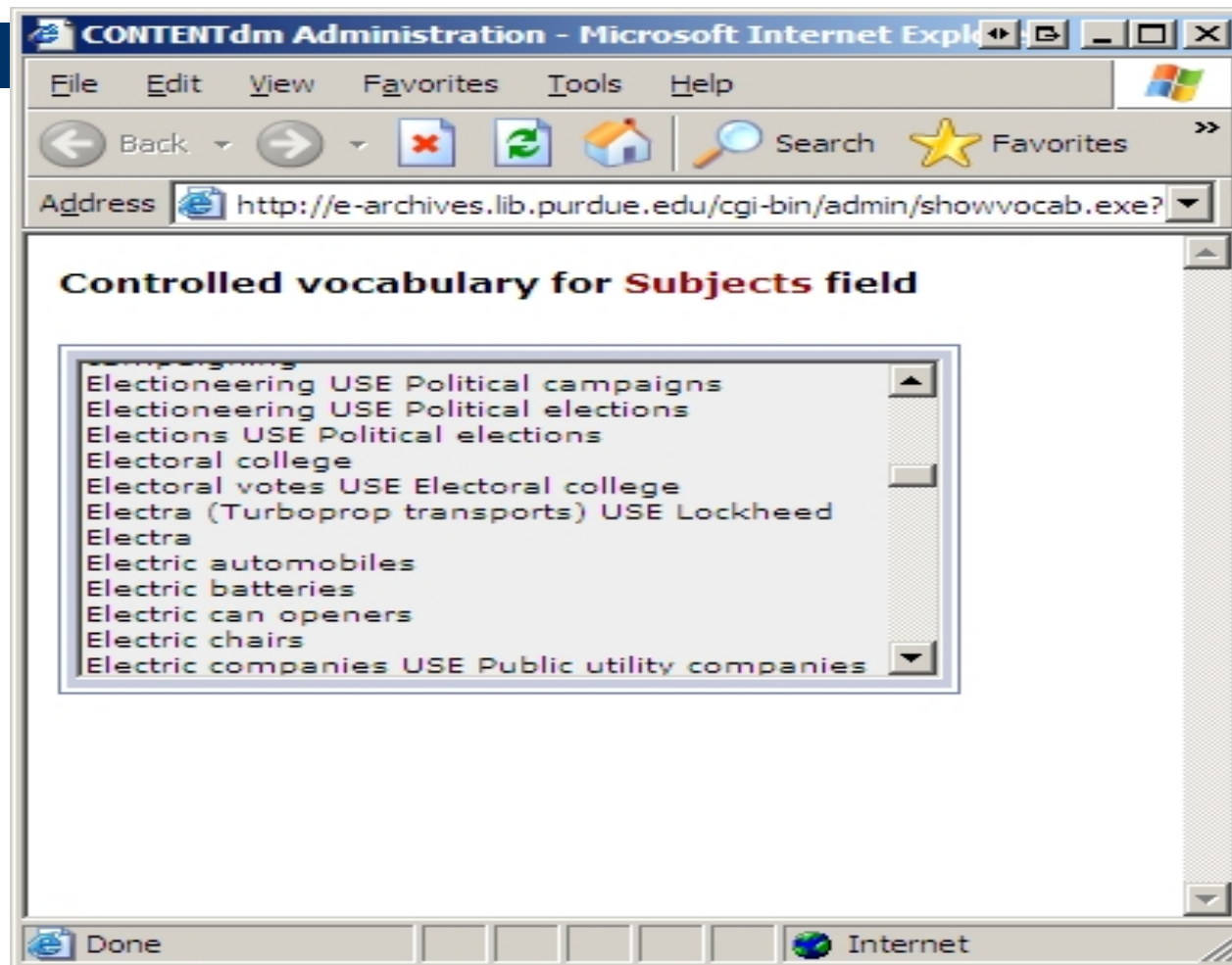
The 'other' vocabulary

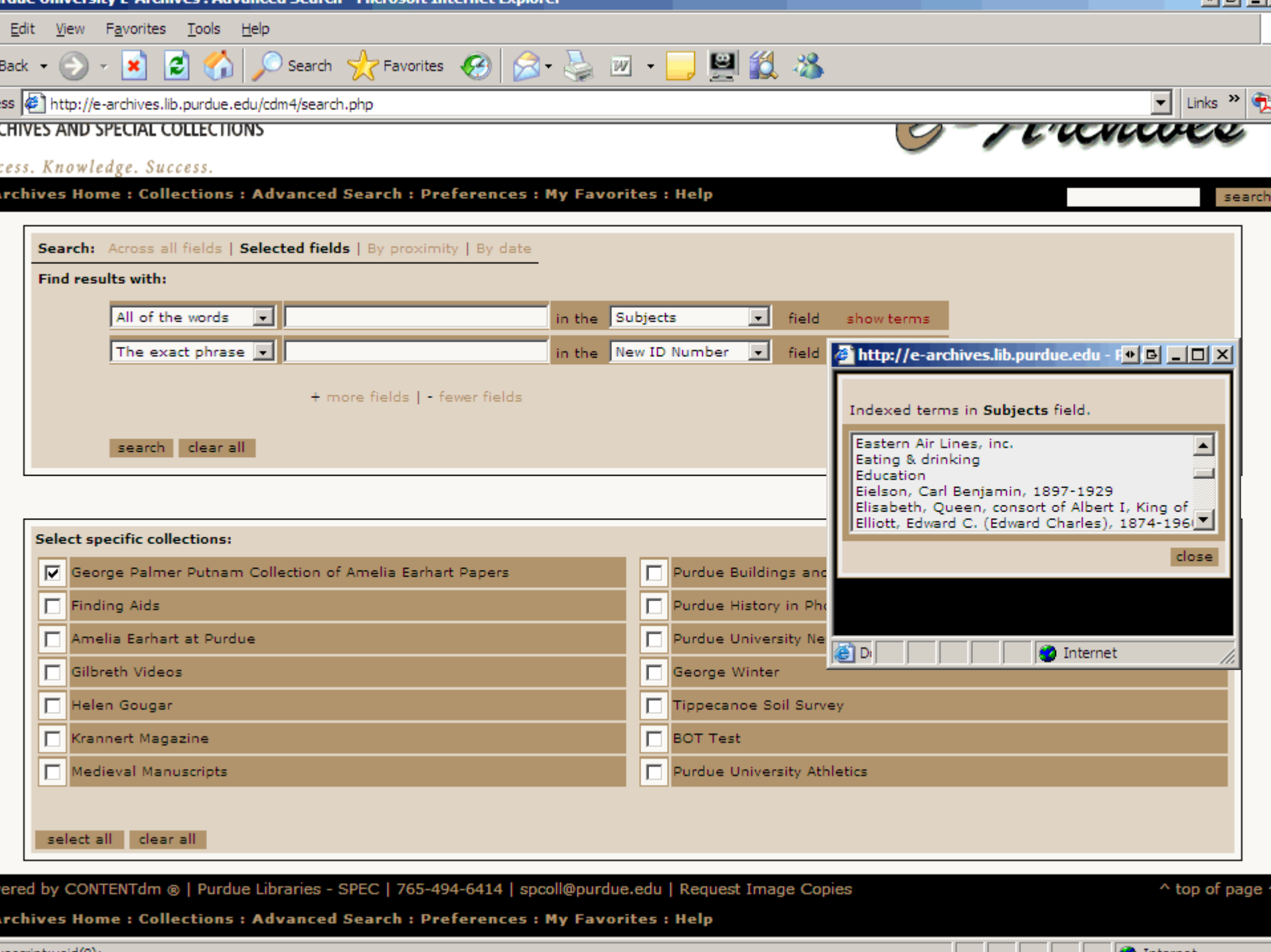
The user's vocabulary

Did you mean: "Bob"



Cross references







| | Field name | DC map | Data type | Large | Search | Hide | Vocab | | add field |
|----|---------------------|--------------|-----------|-------|--------|------|-------|--|---------------|
| 1 | Old ID Number | Identifier | Text | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 2 | New ID Number | Identifier | Text | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 3 | Title | Title | Text | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 4 | Description | Description | Text | Yes | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 5 | Creator | Creator | Text | No | Yes | No | Yes | <input type="button" value="move to"/> | edit delete |
| 6 | Repository | Publisher | Text | No | No | No | No | <input type="button" value="move to"/> | edit delete |
| 7 | Contributor | Contributors | Text | No | Yes | No | Yes | <input type="button" value="move to"/> | edit delete |
| 8 | Date of Original | Date | Date | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 9 | Type | Type | Text | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 10 | Format | Format | Text | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 11 | Extent of Original | Source | Text | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 12 | Orientation | None | Text | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 13 | Language | Language | Text | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 14 | Collection | Relation | Text | No | No | No | No | <input type="button" value="move to"/> | edit delete |
| 15 | Collection Web Site | Relation | Text | No | No | No | No | <input type="button" value="move to"/> | edit delete |
| 16 | Rights | Rights | Text | No | No | No | No | <input type="button" value="move to"/> | edit delete |
| 17 | Subjects | Subject | Text | No | Yes | No | Yes | <input type="button" value="move to"/> | edit delete |
| 18 | Date Scanned | None | Date | No | Yes | No | No | <input type="button" value="move to"/> | edit delete |
| 19 | Capture Device | None | Text | No | No | No | No | <input type="button" value="move to"/> | edit delete |

Decisions on labels: This?

- **Field Name**

LC Subject

TGM Subject

Local Subject

TGM Form or Genre

- **Dublin Core Map**

Subject

Subject

Subject

Subject

Or This?

- **Field Name**

Topics

Forms/Genres

- **Dublin Core Map**

Subject

Subject

Or This?

- **Field Name**

Persons

Corporate Bodies

Topics

Local Headings

Forms and Genres

- **Dublin Core Map**

Subject

Subject

Subject

Subject

Subject

Decision points on desirables

- What metadata format will allow the library to specify which vocabulary any term came from
- What digital asset management system will allow the library to specify which vocabulary any term came from
- What user needs for best search and retrieval
- The “i” word: interoperability

[Home](#)[Search](#)[Help](#)

•

[About](#)[Using OA1ster](#)[News](#)

Search our test portals --
[DLF](#), [MODS](#) -- to use features
we're researching.

Tips:

If you enter more than one word in a search box, the words will be treated as a phrase and searched together, e.g., entering *fancy dress* will find records that have the word *fancy* next to *dress*. No quotations are necessary.

Use an asterisk with a word stem to find variations on the word, e.g., searching for *civ** will match *civ*, *civil*, *civic*, *civilization*, and *civilian*.

Authors may be in inverted order, e.g., if *emily dickinson* doesn't return results, try *dickinson emily*.

See [search help](#) for more tips.

See [issues regarding access](#) to records and digital objects.

Search within fields [\(help\)](#)

Use a word or phrase, e.g., diploma*, fancy dress

 Limit by resource type [\(help\)](#)

Choose one or more

all types
text
image
audio
video
dataset

Choose sorting of results [\(help\)](#)

e.g., results in date descending order, from 2002 to 1999

Standards

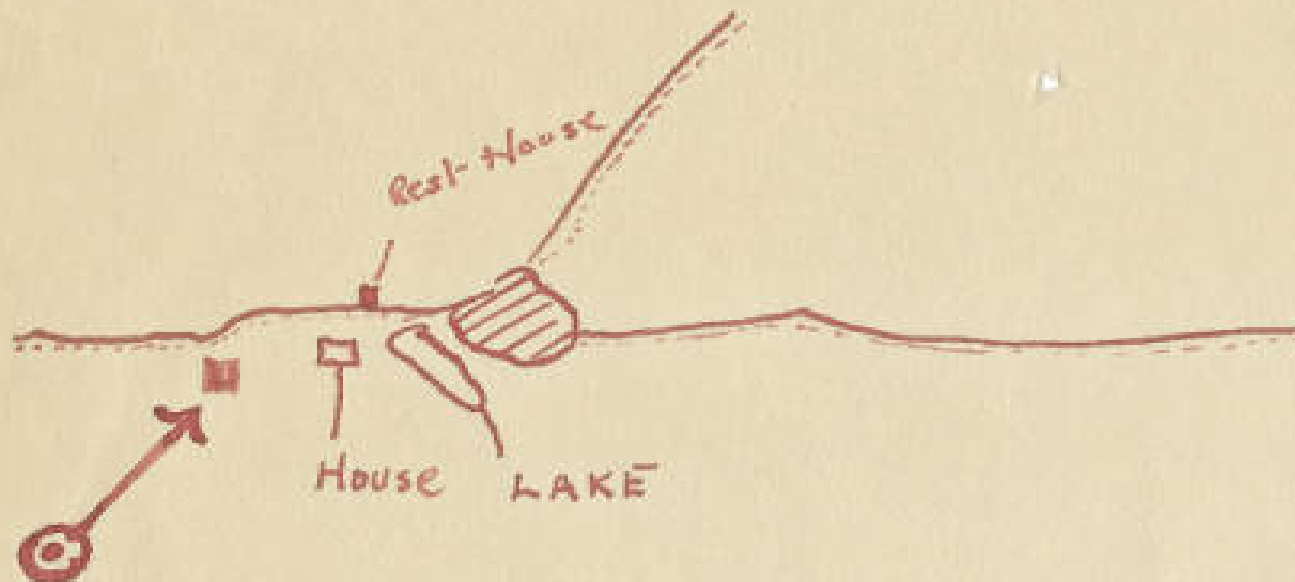
- Dublin Core schema → The structure
- TGM, LCSH, Local subjects → The content
- End result when cross-searching → Bibliographic Babel?

Controlled vocabulary: what we lacked

- Cross-references don't show in the public display
- Unable to search single subject fields across collections
- Simple Dublin Core designed to help resource discovery and to be a mapping tool, but doesn't provide sufficient information for other purposes

LCNAF official name:

Fasher (Sudan)



EL FASHER



Man playing with a leopard cub



Fort-Lamy (Chad) or
N'Djamena (Chad)

| | |
|------------------|--|
| ID Number | XI.B.10 |
| Item ID Number | b11f5i7 |
| Description | Man playing with a leopard cub |
| Repository | Photograph of an unidentified man playing with a leopard cub, Fort Lamy, ca. 1937 [small photograph, 3 copies] |
| Date of Original | Purdue University Libraries, Archives and Special Collections |
| Image | 1937 |
| Format | image |
| | 102 |

Documentation

- Data dictionary useful as a place to document local decisions
- Good enough? For the machine as well as the human indexer?

Dublin Core Usage in Contentdm

Purdue University Libraries

The official Dublin Core page is the Dublin Core Metadata Initiative: <http://www.dublincore.org>

Purdue Libraries will follow best practices as delineated in Dublin Core Metadata Best Practices, CDP Metadata Working Group: <http://www.cdphheritage.org/cdp/documents/CDPDCMBP.pdf>

| Field Name | Dublin Core | Description |
|----------------|-------------|--|
| Title | Title | <p>Two types: (1) transcribed and (2) cataloger-supplied.</p> <p>(1) Transcribed: Type as-is into Title field.</p> <p>(2) Cataloger-supplied: Formulate a succinct title based on the prominent aspects of the resource.</p> <p>The title should strive to uniquely identify the image. For photographs, prefer any written information from the back of the photo. For telegrams, letters, etc. follow formulation as prescribed in AACR2.</p> <p>Omit initial articles.</p> <p>Brackets, quotation marks, and ending punctuation are not used.</p> |
| Creator | Creator | <p>Person/institution chiefly responsible for the creation of the resource. Photographer of a photograph is the creator. Writer of a letter is the creator.</p> |

Digital Initiatives

- [Home](#)
- [About Us](#)
 - [Overview](#)
 - [Team](#)
 - [Toolkit Overview](#)
 - [Hardware](#)
 - [Software](#)
 - [Metadata](#)
 - [Project Manual](#)
- [Digital Collections](#)
- [e-Archives](#)
- [Requesting Images](#)
- Exhibits
 - [George Ade](#)
 - [Amelia Earhart](#)
 - [John Mc Cutcheon](#)
 - [George Winter](#)
- [Related Links](#)
- [Finding Aids](#)

Desirable

- Rich set of descriptive terms in a controlled vocabulary
- Fewer instances of mixed vocabularies without explanation
- If vocabularies are mixed, make that clear via documentation, field labels, etc.
- Cross-reference display for users
- Even better: more flexible metadata models
- Even better: user access to thesaurus for interaction with hierarchy, related terms, broader terms, narrower terms, etc.



Courtesy Purdue University Libraries, Archives and Special Collections



Courtesy Purdue University Libraries, Archives and Special Collections







References

- Baca, Murtha, "Practical issues in applying metadata schemas and controlled vocabularies to cultural heritage information," *Cataloging & classification quarterly*, v. 36, no. 3/4 (2003)
- Best Practices for OAI Data Provider Implementations and Shareable Metadata (wiki, viewed Aug. 27, 2007)
<http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/SubjectPractices>
- *Best practice guidelines for digital collections at University of Maryland Libraries*
<http://www.lib.umd.edu/dcr/publications/best-practice.pdf>

References, continued

- Buckland, Michael, et al., “Mapping entry vocabulary to unfamiliar metadata vocabularies,” *D-Lib magazine*, Jan. 1999.
- Dalmau, Michelle et al., “Integrating thesaurus relationships into search and browse in an online photograph collection,” *Library hi tech*, v. 23, no. 3 (2005)
- Gonzalez, Allegra, “Consistent metadata from disparate sources,” PowerPoint presentation at wu2007
<http://www.consultdm.com/wu2007/presentations/wu2007-allegra-gonzalez.ppt>

References, continued

- Greenberg, Jane, “User comprehension and searching with information retrieval thesauri,” *Cataloging and classification quarterly*, v. 37, no.3/4 (2004)